SECURITY CLASSIFICATION OF THIS PAGE

AD-A238 235

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release: distribution unlimited
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S) AEOSR-TR- 91 0597
6a NAME OF PERFORMING ORGANIZATION Mass. Institute of Technology	6b OFFICE SYMBOL (If applicable)
7a NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Res. N/L	
6c. ADDRESS (City, State, and ZIP Code) Dept. of Brain & Cognitive Sciences 79 Amherst St. E10-120 Cambridge, MA 02139	7b. ADDRESS (City, State, and ZIP Code) Bldg. 410 Bolling Air Force Base Washington, DC 20332-6448
8a NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b OFFICE SYMBOL (If applicable) NL
9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR 89-0504	
8c ADDRESS (City, State, and ZIP Code) Same as 7b	10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO	PROJECT NO	TASK NO	WORK UNIT ACCESSION NO
61102F	2313	A9	

11 TITLE (Include Security Classification)
"Top-Down" Influences on "Bottom-Up" Processing

12 PERSONAL AUTHOR(S)
Whitman Richards

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Annual	FROM 9/89 TO 5/91	6 May 1991	5

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Vision, AI, Cognition, Neurophysiology, Visual Psychophysics, Dynamical Systems

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

Over the past year and one-half, our research effort can be broken down into four categories:

1. A formal framework for percepts.
2. A logic for reasoning about percepts.
3. Experiments related to the above.
4. Seeking chaos in high-level visual processing.

The attached briefly describes progress in these areas.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL Dr. Genevieve Haddad	22b TELEPHONE (Include Area Code) (202) 767-5021
22c OFFICE SYMBOL NL	

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

U.S. Government Printing Office: 1985—507-047

91-04769

91 7 11 063

Interim Report on Top-Down Influences on Bottom-Up Processing (May 1991)

Whitman Richards
MIT E10-120
79 Amherst St.
Cambridge, MA 02139

617-253-5776
whit@ai.mit.edu

AFOSR 89-0504

Over the past year and one-half, our research effort can be broken down into four categories:

1. A formal framework for percepts.
2. A logic for reasoning about percepts.
3. Experiments related to the above.
4. Seeking chaos in high-level visual processing.

1. A Formal Framework for Percepts

We have two thrusts here. The first is to offer a formal definition of a "Percept", and to explore the consequences. The second is concerned with specifying conditions that should hold if an image property is to be a "good feature" - namely one from which reliable inferences can be made.

Many are studying "Perception". So then, just what is a Percept? Without a formal definition, how do we decide whether a particular machine or biological state (or model output) qualifies as a perception? Surprisingly, the first formal definition of a percept was offered only two years ago (Jepson & Richards, 1989, 1991). The insight was to place a partial order upon possible (i.e. legal) interpretations of the image data. (The "snapshot" of any region of the image generally has many possible interpretations.) Within such an order, a percept can then be defined as a maximal node. To create the order it is necessary for the perceiver to choose candidate models of the world (with associated premises), and to test the "goodness of fit" of these models with the observed data. It can be proven that such a hypothesize-and-test approach will generally have several "best-fit" solutions that are locally maximal. The important point, however, is that "top-down" knowledge dictates in part the ordering of the interpretations of the sense data, and hence the resultant percepts. Of course, such "top-down" effects upon our percepts have been known for some time by the experimental psychologists. What's new is that now we have a formal framework that expresses precisely how this knowledge is used.

Our framework raises several formal issues that are currently under study (the experimental questions are listed in the next section):

(i) Given several locally maximal nodes, how can one legally move from one to another, or when are such transitions not allowed? (i.e. when we move from one perceptual state to another, what can and what can not change?)

(ii) What logic can be used to reason about how the models fit the data? (Some kind of default logic is needed – see Section 2 for one new logic that might work.)

(iii) What are the formal relations between our "Lattice Theory for Percepts" and neural net implementations?

(iv) How are models indexed? (Here, we believe we can show that choice of coordinate frame and part-based grouping is critical.)

Related to the above is a second thrust, namely an answer to the question, "What makes a good feature?" (Richards & Jepson, 1991). Certain image properties (such as colinearity, parallel lines) allow strong inferences about the 3D configuration that projects into these image features. Other image projections, such as a "T" do not generally support strong inferences, because they can arise from many different kinds of events, such as two twigs abutting, a stick on a surface (like a table leg), or the occlusion of one surface by another. Given a particular model world, we can show how to enumerate all features in the image that support strong inductions. In other words, we can specify which image properties are worth looking at, and what they are likely to "mean". Many of these same features also provide useful indices to classes of models.

2. A New Logic

Little is known about how we "reason" about a collection of image entities. Certainly we have heuristics that suggest certain groups of image features "belong together", etc. However assigning likelihoods to various groupings is not reasoning, but only information that the reasoning process can use (Pearl, 1988). Recently, Bennett & Hoffman (1990) proposed a new "Lebesgue" logic that looks attractive for perceptual reasoning, provided one makes a simple change. Rather than using continuous probability measures for events, we suggest using the rank order of the event measures. This leads to a modification called "Order Logic". By early next year we hope to have shown that "Order Logic" is a variety of default logic that will support a perceptual reasoning process that uses fallible premises.

3. Experiments

A simple application of our "Lattice Theory" is resolving conflicts between image interpretations offered by the different vision modules, such as stereo, structure-from-motion, shape, etc. Generally the outputs will not agree. (A very common example is when you watch TV. Your stereo disparity world is flat, yet your structure-from-motion module easily recovers 3D shape. Which is "correct"?) We have explored point and line displays, such as the Ames Trapezoid window plus bar, in order to show how the observer's chosen interpretation "makes sense", given certain premises (hypotheses) about the world which resides in his knowledge base. The trick, then is to discover these premises.

So far, our most striking finding is that the perceiver first assigns a 3D coordinate frame to the image and to groups of image entities that are "part-like". (We do not have a formal scheme yet that predicts which entities will be the "parts".) The assigned coordinate frame is a "guess" and critical, and hence is a "top-down" premise that guides further image interpretation. This same start-up effect (specifically interpreting "blocks-world" figures and in assigning figure-ground), and can lead to "garden path" percepts. The example of this that we are studying is a rigid configuration in motion that appears non-rigid.

We have also examined in detail two blocks-world examples, again with the aim of discovering the premises used by most when building interpretations. Support under gravity, attachment of parallel faces, colinearity of aligned edges or faces, together with the above coordinate frame premise, are typical examples of what we commonly find. We are also exploring the model parameterizations people use, and how these force certain categorical perceptions (Feldman).

On a completely different tack, we began some parametric studies to determine whether the switching properties for multistable percepts can be deduced. (These studies would help us understand the dynamics underlying movement from one percept to another in a lattice of allowable percepts.) Because time is a parameter invariant across mechanisms, we have chosen to determine the temporal properties of "blocking" or "switching" between states (nodes) in a lattice of partial orderings. To date our best data come from multistable percepts such as the crater illusion, or conflicting cognitive contours. However our understanding of the "switch" is still incomplete (see below).

4. Chaos and Dynamical Systems

The above "switching" studies show that multi-stable percepts entail non-linear dynamics and are definitely not Poisson processes. There is a strong hint of an underlying chaotic mechanism behavior of dimension roughly 3.8. However we have not yet been able to translate our data into an attractive (!) chaotic model. To date only the rough shape of the phase space has been determined.

5. Publications

What Is a Percept? (Jepson & Richards) Cognitive Science Occasional Paper #43, Center for Cognitive Science, MIT, Cambridge, MA 02139.

A Lattice Theory for Percepts. (Jepson & Richards) Submitted to *Perception* 5/91.

Integrating Vision Modules. (Jepson & Richards) To appear in IEEE-SMC.

Self-Calibrated Collinearity Detector. (Moses, Schecklman & Ullman) *Biol. Cyb.*, **63**:463 - 475 (1990).

Curvature and Separation Discrimination at Texture Boundaries. (Wilson & Richards) *Jrl. Opt. Soc. Am.*, under review.

In preparation:

What Makes a Good Feature? (Richards & Jepson). [Cornell presentation, June 1991.]

Why Is Snow So Bright? (Koenderink & Richards).

Reasoning Under Uncertainty: Lebesgue Logic and Order Logic. (Bennett, Hoffman & Richards). [*Cog. Sci. Proc.*, Aug. 1991.]

Choosing a Coordinate Frame. (Richards & Subirana). [See "What Is Figure?" ARVO 1991, for brief presentation.]

Talks (Symposia):

Perception and Perceivers. (Vision and 3D Representation, Univ. Minn., May 1989).

What Is A Perception? (Assimilation in Man and Machines, Univ. Michigan, June 1990).

Perception, Computation and Categorization. (Cog. Science Soc., July 1990).

Integrating Vision Modules. (Spatial Vision Conf., York, June 1991).

What's A Good Feature? (Neural Substrates of Perceptions, Cornell, June 1991).

6. Funds

We anticipate roughly a $1000 shortfall at the end of the first 25 months of the grant (i.e. 30 Sept. 1991). However, Shimon Ullman will be at Weizman for the academic year 1991-92, returning to MIT for two one-month periods, for which he would be paid. This will save us roughly $10,000 over the year, including overhead. We would like to use this to repair our Sun, and to acquire a second Mac III for experiments.